ANALYSIS OF THE RELIABILITY OF THE COGNITIVE ABILITY TEST AS A PREDICTOR OF INTELLECTUAL GIFTEDNESS

A Field Report

Presented to

The School of Graduate Studies

Drake University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Education

---

by

Mona Ward O'Brien

May 1987

1987
.06

ANALYSIS OF THE RELIABILITY OF THE COGNITIVE

ABILITY TEST AS A PREDICTOR OF

INTELLECTUAL GIFTEDNESS

by

Mona Ward O'Brien

Approved by Committee:

Paul Joslin

Dr. Paul Joslin, Chair

Hilda Lee Williams

Dr. Hilda Williams

George S. Lair

Dr. George Lair
Dean of the Graduate School of Education

ANALYSIS OF THE RELIABILITY OF THE COGNITIVE ABILITY TEST AS A PREDICTOR OF INTELLECTUAL GIFTEDNESS

An abstract of a Field Report by
Mona Ward O'Brien
May 1987
Drake University

The problem. This study examined the reliability of the Cognitive Ability Test by Thorndike and Hagen as a tool for identification of gifted students.

Procedure. Statistical analysis using a chi-square and a t-test were used to examine the reliability/ repeatability of this test. The first grade scores were compared with the fourth grade scores of the same students using a correlated t-test. Within group comparisons were based on a chi-square analysis.

Findings. This study showed that there was a high relationship between Test 1 and Test 2, therefore, Test 1 is highly predictable of Test 2. However, this was not true in the upper range scores (those in the 130 and above range).

Conclusions. This study found that the upper range scores in the Cognitive Ability Tests do not repeat in a subsequent test of the same students.

Recommendations. Use of the Cognitive Ability Test as a predictor of gifted and talented students should be reevaluated since this study found that the very range where the scores are used for classification in the gifted and talented programs is the range where the test lacks repeatability, and thus reliability.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Statement of the Problem

Most school districts recognize the special needs of the gifted and talented students and attempt to provide programs for this special population. Among educators there is little agreement as to what exactly constitutes a talented and gifted student. There does not exist a clear, generally agreed upon definition of what gifted and talented is.[1] Outstanding, or extraordinary, or above-average ability is assumed to be a characteristic of giftedness but how does a school identify outstanding ability in young students? What indicator can be used? What results can be measured quantitatively in order to provide accountability to the community? Since talented and gifted programs are for a few students, a method of screening must be used to limit admittance. What factors can be used to identify students for gifted and talented programs?

Several methods are available for quantitative identification of high potential or gifted students. These

---

[1] Francoys Gagnel, "Giftedness and Talent: Reexamining of the Definitions," Gifted Child Quarterly 29 (Summer 1985): 103-11.

methods include scores on intelligence type (cognitive ability) creativity tests, grades, nominations, and scales using checklists. Of these, the most commonly employed is some form of the intelligence or cognitive ability test. The results of a national survey by Alvino, McDonnel, and Richert on identification of gifted and talented revealed that by far the single, most often used means of identification was the group intelligence test, or today's equivalent, which is labeled cognitive ability, school achievement, or scholastic aptitude test.[1] An interview with Dr. Lee Wolf of Iowa's Department of Education revealed that at least 90 percent of Iowa's school districts that have gifted and talented programs employ some group I.Q. type test for determining admittance.[2] Thus, as varied as the researchers and tools of identification are, the fact remains that I.Q. type tests (or today's replacements called Cognitive Ability, school, or scholastic aptitude tests) are the most commonly used tool in identification of gifted students.

Purpose

The purpose of this study is to analyze the consistency of the Cognitive Ability Test (CAT) as a tool for

---

[1] James Alvino, Rebecca McDonnel, and Susanne Richert, "National Survey of Identification Practices in Gifted and Talented Education," Exceptional Children 48 (Oct. 1981): 124-32.

[2] Lee Wolf, telephone interview, 21 July 1986.

identification of gifted and talented students. The scores of 1,352 students who had taken the Cognitive Ability Test more than once were examined. By comparison of dual scores on the same child, valuable information as to the repeatability of group test scores was gathered. Analysis of this information may be useful for those in gifted education in determining if group I.Q. type tests, specifically the CAT, are reliable indicators to be used for admission into gifted and talented programs. Three characteristics of a sound test, as established by the Standards of Educational Psychological Testing are "adequate standardization, reliability, and validity."[1]

Research Hypothesis and Prediction

The hypothesis for this study is that the upper range scores obtained in the first grade test would not be repeated on a reliable basis in the fourth grade (or second) test of the same student. Statistically, some regression toward the mean would be expected. However, differentiation at the upper range may be lacking in a group test. Therefore, there will be a higher than expected regression indicating a lack of consistency of the Cognitive Ability Test as a predictor of upper range scores on subsequent repetition.

---

[1] Alan S. Kaufman and Patti L. Harrison, "Intelligence Tests and Gifted Assessment: What Are the Positives?" Roeper Review 8 (Feb. 1986): 155.

Operational Definitions

The following operational definitions are used in this study:

Test 1: Cognitive Abilities Test, Form 3 Primary Battery by Thorndike and Hagen. The single raw score was converted to a Standard Age Score (SAS).

Test 2: Cognitive Abilities Test, Multilevel edition Form 3 Level B by Thorndike and Hagen. The three scores that result are converted to a Standard Age Score and averaged for a single score for comparison.

Categories: The scores were divided into seven categories as follows:

category 1 = scores 140 and above (test tops at 150)

category 2 = scores 130-139

category 3 = scores 120-129

category 4 = scores 110-119

category 5 = scores 100-119

category 6 = scores 90- 99

category 7 = scores 89 and below

Cells were created at ten point intervals since scores were to be interpreted as ten point bands to allow for error of measurement. Thus the cells are as follows:

cell 1 = loss of 20 or more points

cell 2 = loss of 11-19 points

cell 3 = loss or gain of 10 points

cell 4 = gain of 11-19 points

cell 5 = gain of 20 or more points

Gains and losses refer to changes between test scores at the first and fourth grades.

Limitations

The students selected for this study were present in this district in both their first and the fourth grade years and took both tests. Parental permission was requested and a few students in the school district who were present in both the first and the fourth grade years did not take the Cognitive Ability Test because the parent declined to grant permission for them to do so (the number was small).

Selection of Sample

The population for this study was all students in a suburban population who were enrolled in that district from 1979 to 1983 for the first grade test, and were still enrolled in 1982 to 1986 for the fourth grade test. Five consecutive years of dual scores were recorded for this study.

CHAPTER 2

Review of Liberature

Among educators there is little agreement as to what exactly constitutes a gifted and talented student. One of the most famous longitudinal studies on gifted and talented students is Lewis Terman's Genetic Studies of Genius. Terman's definition of giftedness was restricted to those who scored in "the top 1% in general intellectual ability, as measured by the Stanford-Binet Intelligence Scale or a comparable instrument."[1]

An expansion of intelligence as the sole criteria for gifted and talented is described in the Marland report of 1972.

> Gifted and talented children are those identified by professionally qualified persons who by virtue of outstanding abilities are capable of high performance. . . . Children capable of high performance include those who have demonstrated any or potential ability in any of the following areas, singly or in combination:
> 1. general intellectual ability
> 2. specific academic aptitude
> 3. creative or productive thinking
> 4. leadership ability
> 5. visual and performing arts aptitude
> 6. psychomotor ability.[2]

---

[1] L. M. Terman, Genetic Studies of Genius: Mental and Physical Traits of a Thousand Gifted Children (Calif: Stanford Univ. Press, 1926), 43.

[2] S. P. Marland, Education of the Gifted and Talented. Report to the Congress of the United States by the U.S.

E. P. Torrance advocates recognition of creativity as a component of giftedness and thus use of a test of creativity for selection.[1] Because creativity is an identifiable skill not measured by tests which are based on questions with one right answer, group intelligence type tests may fail to identify the creatively gifted and talented student.

J. Renzulli and Marcia Delcourt combined creativity with above average intelligence and task commitment to further broaden the definition of what constitutes gifted and talented. A challenge to this theory was issued by J. R. Whitmore since it eliminates what she terms the gifted underachievers who are intelligent but lack motivation or commitment. Abraham Tannenbaum considers intelligence but stresses special abilities and inter- and intrapersonal relationships. Howard Gardner espouses a theory of multiple intelligences in which there are at least six distinct intelligences: linguistic, logical-mathematical, spatial, musical, bodily-inesthetic, and personal. Gardner points out that what is valued varies with cultures.[2]

Commissioner of Education, Vol. 1 (Washington, D.C.: U.S. Government Printing Office, Aug. 1971), ix.

[1] E. P. Torrance, "The Role of Creativity in Identification of the Gifted and Talented," Gifted Child Quarterly 28 (1984): 153-56.

[2] Joseph Renzulli and Marcia Delcourt, "The Legacy and Logic of Research on the Identification of Gifted Persons," Gifted Child Quarterly 30 (Winter 1986): 20-23; J. R.

R. S. Sternberg favors a triarchic approach to intelligence. In his view, today's intelligence type tests measure logical thinking but not the way one thinks, or synthetic abilities, and the present tests favor precocity which is not necessarily giftedness. A shortcoming of the intelligence type tests today is evidenced by the fact that many whose accomplishments have resulted in their being viewed as gifted and or talented would not be identified by these commonly used tests (Lee Iacocca, Martin Luther King, Jr., Albert Einstein, and Ernest Hemingway to name a few).[1]

The origin of the intelligence test lies with Alfred Binet's attempt to devise a test to identify retarded children whose needs would not be met in public classroom. The group use of intelligence tests took a leap into respectability in the U.S. when employed by the Army during World War I. After that the group intelligence test found acceptance in the public schools. Intelligence tests fell out of favor in the 60s and 70s; they were found to favor and, conversely, exclude, certain groups. (The problem of bias is well documented and will not be addressed in this paper.)

Whitmore, Giftedness, Conflict, and Underachievement (Boston: Allyn & Bacon, 1980); Abraham Tannenbaum, Gifted Children (New York: Macmillan, 1983), p. 27; Howard Gardner, Frames of Mind (New York: Basic Books, 1983), p. 25.

[1] Robert J. Sternberg, "Identifying the Gifted Through IQ: Why a Little Bit of Knowledge is a Dangerous Thing," Roeper Review 8 (Feb. 1986): 143-47.

There is a lack of agreement as to the reliability of the intelligence test. The correlation between the WISC and the Stanford Binet is .8 on the verbal and 6.5 on performance; however, these are both individual rather than group tests.[1]

The claims to repeated high correlation of I.Q. scores after age seven is based on individual intelligence test results, not group tests.[2] Pegnato and Birch found that group I.Q. tests do not correlate with individual I.Q. tests.[3] When using a 130 cut off range on a group I.Q. test, only 22 percent of whose who scored over 135 on the Stanford Binet were found. This study was repeated by Martinson and Lessinger using over 300 subjects.[4] The results at the upper range of the group I.Q. test (cut off of 130) was again unreliable in that less than one-half of those who scored higher than 130 on the revised Stanford Binet were in the identified group. Group intelligence test results are not interchangeable with individual intelligence test results.

---

[1] Evelyn Sharp, The IQ Cult (Toronto: Longman Canada Ltd., 1972), 64.

[2] Kaufman and Harrison, p. 155.

[3] C. W. Pegnato and J. W. Birch, "Locating Gifted Children in Junior High Schools Comparison Methods," Exceptional Children 25 (1959): 300-04.

[4] R. A. Martinson and L. M. Lessinger, "Problems in Identification of Intellectually Gifted Pupils," Exceptional Children 26 (1960): 227-42.

Group intelligence tests do correlate with school success, but not job proficiency.[1] Since school I.Q. tests are like school tests, that a correlation exists is not surprising. Both tests use questions based on one right answer and both use present acquired knowledge. Because of the nature of a group test, how the student thinks through or solves a problem is never determined.

In summary, a review of the literature revealed a lack of a precise, agreed-upon definition of what is gifted and talented. Lacking an agreed-upon definition of gifted and talented, the intelligence test falls short on validity as a characteristic of a good test of gifted and talented according to the Standards of Educational and Psychological Testing. Since we aren't able to define intelligence, it probably can't be measured in a single score. It also seemed evident from the literature that there is little support from those in research and practitioners for using group intelligence tests.

Today, a group I.Q. test cannot be used to identify slow or poor learners for placement into remedial or special classes. It is ironic that as the use of lower range scores of the group I.Q. type test fell into disfavor as a reliable tool for placement into special classes, the

---

[1] Nancy M. Robinson and Diane L. Chamrad, "Appropriate Uses of Intelligence Tests with Gifted Children," Roeper Review 8 (Feb. 1986): 161.

use of the upper range scores of the very same tests is being widely utilized as a means for identification and thus placement into special classes for gifted and talented students.

There is a lack of agreement among researchers and educators of the gifted and talented as to a definition. There is a lack of agreement among researchers and educators of the gifted and talented as to the best tool for identification. "There is often a lag between what is known and what is practiced. . . . Probably no area in education would illustrate the truth of this adage more profoundly than gifted education, especially in matters pertaining to definition and identification."[1] Gardner points out in his survey of studies that intelligence is what is valued by a society and therefore varies from culture to culture.[2] Despite this lack of agreement on a definition of gifted and talented, the group I.Q. type test is the most used means of identification for gifted and talented.

As Sternberg points out, the I.Q. type test favors the quick problem solver using conventional methods and past acquired knowledge.[3] As the studies have shown, I.Q.

---

[1] Donald J. Treffinger and Joseph Renqulli, "Giftedness as Potential for Creative Productivity: Transcending IQ Scores," Roeper Review 8 (Feb. 1986): 151.

[2] Gardner.

[3] Sternberg, "Identifying the Gifted."

type tests correlate with school achievement but not with job proficiency. Sternberg and others call attention to the fact that precocity and conventional thinking is not what we assign gifted or talented labels to in adults. Although the literature fails to show a correlation between I.Q. type tests and future achievement, this type of test is used extensively. If, as this study shows, group tests lack repeatability in the very range of scores used for gifted and talented identification, then perhaps the use of these tests needs to be reevaluated.

CHAPTER 3

Analysis of the Data

Procedure

This study analyzes the use of the Cognitive Abilities Test by Thorndike and Hagen as a tool for identification of students for placement into a gifted and talented program. The problem addressed by this study is whether the high scores obtained in the first grade Cognitive Ability Test Form 3, Primary Battery were consistently repeated in the fourth grade year in the Form 3 Level B of the Cognitive Ability Test. In order to check the consistancy/reliability of these findings, a five year study was done. Only students who took the test in first grade and remained to be retested in fourth grade were included in this study.

The Cognitive Abilities Test evolved from the Lorge-Thorndike Intelligence Test.[1] It is designed to measure scholastic aptitude and abstract reasoning. The raw score test results can be converted, via tables, to a universal score which is used in combination with the chronological age of the student to arrive at a Standard Age Score (SAS). The mean for the SAS is 100 with a

---

[1] L. J. Cronbach, Essentials of Psychological Testing (New York: Harper and Row, 1970), 137.

standard deviation of 16. This is similar to the means and standard deviation of the Lorge-Thorndike I.Q. Test.

The Cognitive Abilities Test by Thorndike and Hagen represents contemporary tests which rarely use the word I.Q. Words such as "cognitive ability" and "school" or "scholastic aptitude" have replaced "intelligence tests." However, "at least part of the acceptance of these newer tests depends on whether they are highly similar to traditional intelligence tests."[1] Just as the Lorge-Thorndike Intelligence Test has been replaced by the Cognitive Abilities Test, the Otis-Lennon Mental Ability Test has become the Otis-Lennon School Ability Test.[2] Both of these abilities tests are used in Iowa and elsewhere for the identification of gifted students. Both are group tests with results that are considered highly reliable according to references such as the <u>9th Mental Measurement Yearbook</u>.[3]

This study analyzed the reliability, or repeatability, of scores of students who took the Cognitive Abilities Test Form 3 Primary Battery in the first grade and Form 3 Level B in their fourth grade year. The primary battery yields a single score from a 76-item test. The multilevel

---

[1] Cronbach, p. 135.

[2] Cronbach, p. 137.

[3] James V. Mitchel, Jr., ed., <u>9th Mental Measurement Yearbook</u> (Lincoln: Univ. of Nebraska Press, 1985), 351-52.

edition, Level B yields three scores which are averaged to provide a single score for comparison.

Population

The subjects chosen for this analysis were 1,352 students in a suburban school district. The students of this district were enrolled in eight elementary schools. The standard age scores (SAS) were grouped into seven groups with a ten point spread of scores. The manual provided with the primary battery of the Cognitive Abilities Test advises that scores be interpreted as ten point bands rather than a single score.[1] Therefore, if a student obtains a score of 145, it should be viewed as a score between 140 and 150, thus allowing for error of measurement.

Results

The first grade scores were recorded and then compared to scores from the fourth grade test. A visual analysis of the data in a scattergram (Figure 1) shows a positive linear relationship between Test 1 and Test 2. The fourth grade test is on the y-axis and the first grade test is on the x-axis. The correlation of the two tests is 0.64803. The standard error of estimate is 0.09632. The histogram (see Figure 2) portrays a normal curve for the data.

---

[1] Robert L. Thorndike and Elizabeth Hagen, Cognitive Abilities Test: Examiners Manual, Form 3 (Atlanta: Houghton Mifflin, 1979).

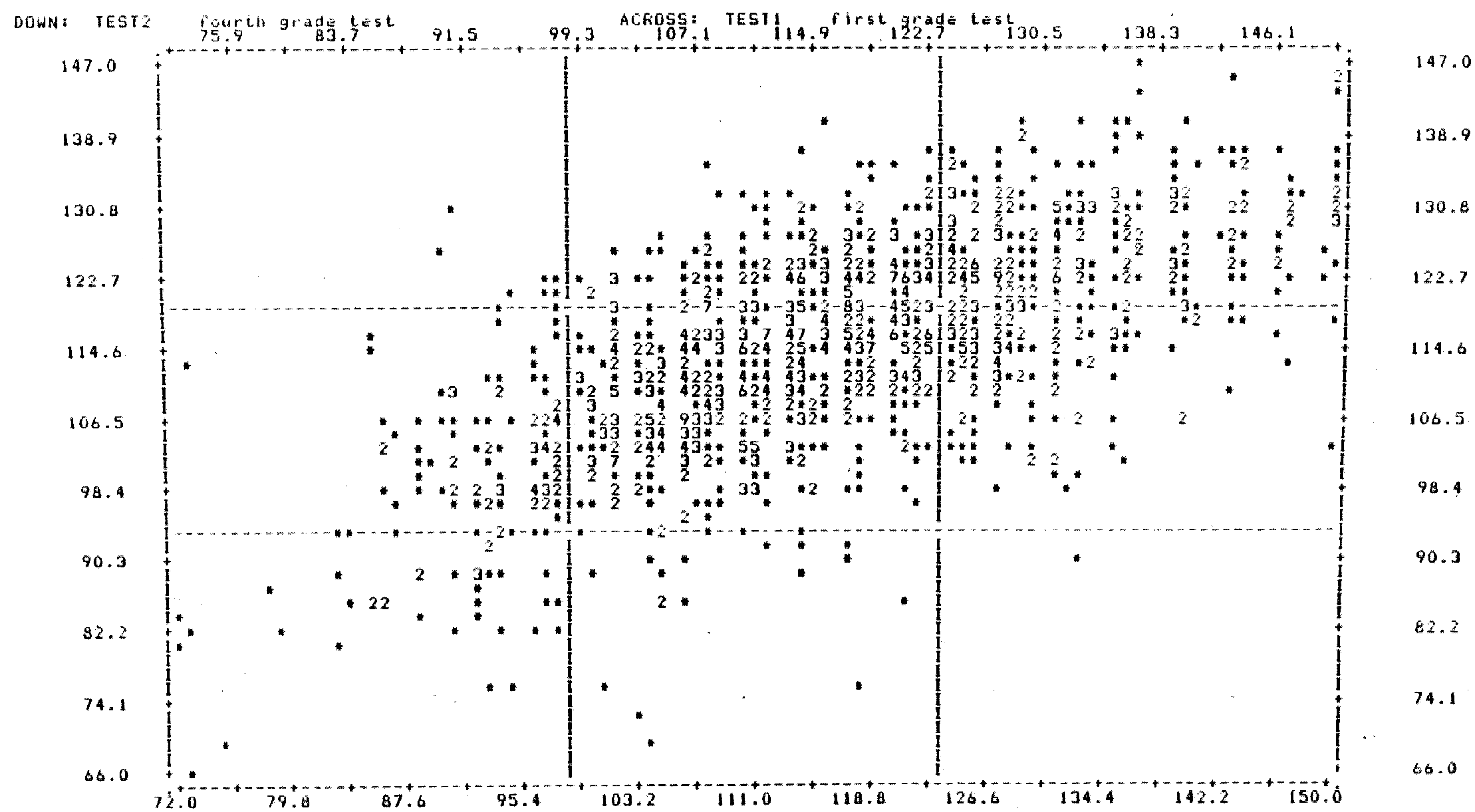

Figure 1

Scattergram

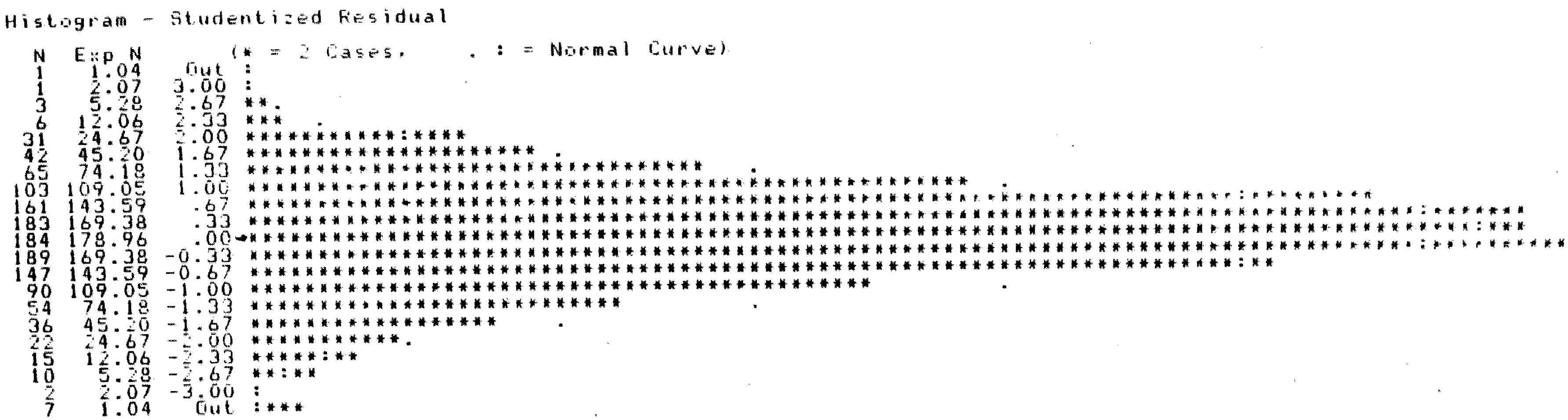

Figure 2

Histogram

A correlated t-test was run and found significant differences on scores in Test 1 and Test 2 (p 0.0000). The correlation ran between Test 1 and Test 2 resulted in a 9.8 error of estimate. The t-test variables are shown in Table 1.

Table 1

t-test Variables

Variable	Test 1	Test 2
Mean	116.12	114.58
Standard deviation	14.285	11.939
Standard error	0.388	0.325
Diff. mean	1.548	
Standard deviation	11.205	
Standard error	0.305	
2-tail correlation	0.648	
Significance	0.000	
t-value	5.08	
Degrees of freedom	1351	
Probability	0.000	

A chi-square analysis was completed to determine in which bands of scores the differences occur, by year and overall. The result was a chi-square of 500.33 with 24 degrees of freedom, eight of thirty-five cells had expected frequencies of less than five. Eight was acceptable because of the large number in this population and because collapsing of the cells would not allow as close an analysis as is obtained with the ten point bands. Although there is an overall relationship between Test 1 and Test 2 and thus a high degree of predictability, this is not so at the upper range or band of scores.

The correlation indicates that the amount of variance in Test score 2 as predicted by Test score 1 is 42 percent (0.4199). When the means are so similar, an explanation for the variance is needed. The regression of Test 1 on Test 2 resulted in a correlation of 0.648, a regression constant of 51.67 and a regression coefficient of 0.541.

Both Test 1 and Test 2 show standard bell curves and similar means so a chi-square test was used to determine if there was a significant difference between the observed frequencies and the expected frequencies. The problem as shown through the chi-square was in the predictive ability of Test 1. Based on the predictability of Test 1, if a student scored 150 on Test 1 (the highes possible) that student would be expected to get 132.67 on the second test, with 68 percent probability that the score would fall between 143 and 123. Similarly, if the Test 1 obtained score was 140, the predicted score would be (51.67 + .54(140) = 127.27 [plus or minus 9.0932]). Since only 3 percent of the student population is admitted into a gifted program (to comply with state guidelines), the students whom the first grade test identified are not, with any predictability, going to reappear in the top category.

Table 2 shows the total five-year change (cells 1-5 on x axis indicate gains or losses) from Test 1 scores as compared to Test 2 scores in each of the seven, ten-point categories. Since the highest score obtainable is 150 there would be no students in Cell 5 of group 1 and few in Cell 4. Likewise, because special classes provided for students scoring in the lower range of scores usually eliminates these students from the average classroom, a zero finding is expected in Cell 1 of Category 7.

The difference between the observed frequency of student scores of 140 and above losing 20 or more points on the second test is 31 in contrast to the expected 5. Thirty-three students in this range lost 11 to 19 points, where only 11 were expected to have such a loss. Again at the 130-139 range of scores, 27 students lost 20 or more points with only 10 expected to lose that many points. Likewise, 43 students in this range lost 11 to 19 points, while only 21 would be expected. This difference between

Table 2

Summary of Five Years of Dual Scores
Population = 1,352

Test Range	1 Lose 20+	2 Lose 19-11	3 Lose 10 or gain 10	4 Gain 11-19	5 Gain 20+	Total
1 150+ 140+	31* 5** 38.8%***	33 10.7 41.3%	16 54.5 20.0%	0 8 0.0%	0 1.9 0.0%	80 5.9%
2 139- 130	27 9.6 17.4%	43 20.6 27.4%	85 105.6 54.8%	0 15.5 0.0%	0 3.7 0.0%	155 11.5%
3 129- 120	13 19.9 4.0%	58 42.7 18.1%	241 218.7 75.1%	9 32.1 2.8%	0 7.6 0.0%	321 23.7%
4 119- 110	6 21.7 1.7%	26 46.6 7.4%	276 238.4 78.9%	38 34.9 10.9%	4 8.3 1.1%	350 25.9%
5 109- 100	7 17.4 2.5%	13 37.3 4.6%	199 190.7 71.1%	50 28.0 17.9%	11 6.6 3.9%	280 20.7%
6 99- 90	0 8.2 0.0%	7 17.6 5.3%	84 89.9 63.6%	28 13.2 21.2%	13 3.1 9.8%	132 9.8%
7 89 and below	0 2.1 0.0%	0 4.5 0.0%	20 23.2 58.8%	10 3.4 29.4%	4 0.8 11.8%	34 2.5%

* actual number
** expected value
*** row percentage

expected and actual does not exist in the normal or mid-range scores, 110-119 and 120-129.

Charts showing each of the five year's actual changes, expected values, and the percentage the changes reflect are included (see Tables 3, 4, 5, 6, and 7). These are included so that researchers will realize that the significance in the five-year summary chart is consistently visible in the yearly charts. Each year in this five-year study, the scores in the 140 and above range have over a 50 percent loss of more than 11 points, more than the ten point band that the manual suggests. Every year the 140 and above range consists of over 3 percent of the total population, therefore, this is the group that would qualify for placement or identification as gifted and talented.

Conclusions

The results of this study show a lack of repeatability in the upper range scores (140+). Although the test booklets state that each score should be viewed as a ten point band, the loss at the upper limits (140+) showed an 11 point or more loss in over 50 percent of the cases studied each and every year. In four of the years the loss of 11 or more points occurred over 75 percent of the time. Upper scores (140+) are the scores that would be used to qualify a student for inclusion into a gifted program. Yet these very scores are the least stable.

Table 3

1982 I.Q. Test
Population = 335

Test Range	1 Lose 20+	2 Lose 19-11	3 Lose 10 or gain 10	4 Gain 11-19	5 Gain 20+	Total
1	7*	11	2	0	0	20
150+	1.4**	3.2	13.3	1.6	0.4	6.0%
140+	35.0%***	55.0%	10.0%	0.0%	0.0%	
2	9	7	25	0	0	41
139-	2.9	6.6	27.3	3.3	0.9	12.2%
130	22.0%	17.0%	61.0%	0.0%	0.0%	
3	4	21	49	2	0	76
129-	5.4	12.3	50.6	6.1	1.6	22.7%
120	5.3%	27.6%	64.5%	2.6%	0.0%	
4	1	8	69	9	0	87
119-	6.2	14.0	57.9	7.0	1.8	26.0%
110	1.1%	9.2%	79.3%	10.3%	0.0%	
5	3	4	41	6	2	56
109-	4.0	9.0	37.3	4.5	1.2	16.7%
100	5.4%	7.1%	73.2%	10.7%	3.6%	
6	0	3	32	6	3	44
99-	3.2	7.1	29.3	3.5	0.9	13.1%
90	0.0%	6.8%	72.7%	13.6%	6.8%	
7	0	0	5	4	2	11
89 and	0.8	1.8	7.3	0.9	0.2	3.3%
below	0.0%	0.0%	45.5%	36.4%	18.2%	

* actual number
** expected value
*** row percentage

Table 4

1983 I.Q. Test
Population = 259

Test Range	1 Lose 20+	2 Lose 19-11	3 Lose 10 or gain 10	4 Gain 11-19	5 Gain 20+	Total
1	9*	9	3	0	0	21
150+	1.4**	3.1	14.8	1.5	0.3	8.1%
140	42.9%***	42.9%	14.3%	0.0%	0.0%	
2	6	9	20	0	0	35
139-	2.3	5.1	24.6	2.4	0.5	13.5%
130	17.1%	25.7%	57.1%	0.0%	0.0%	
3	2	9	48	2	0	61
129-	4.0	8.9	42.9	4.2	0.9	23.6%
120	3.3%	14.8%	78.7%	3.3%	0.0%	
4	0	9	54	5	1	69
119-	4.5	10.1	48.5	4.8	1.1	26.6%
110	0.0%	13.0%	78.3%	7.2%	1.4%	
5	0	2	41	8	2	53
109-	3.5	7.8	37.2	3.7	0.8	20.5%
100	0.0%	3.8%	77.4%	15.1%	3.8%	
6	0	0	13	2	1	16
99-	1.1	2.3	11.2	1.1	0.2	6.2%
90	0.0%	0.0%	81.3%	12.5%	6.3%	
7	0	0	3	1	0	4
89 and	0.3	0.6	2.8	0.3	0.1	1.5%
below	0.0%	0.0%	2.8%	0.3%	0.1%	

* actual number
** expected value
*** row percentage

Table 5

1984 I.Q. Test
Population = 273

Test Range	1 Lose 20+	2 Lose 19-11	3 Lose 10 or gain 10	4 Gain 11-19	5 Gain 20+	Total
1 150+ 140	6* 1.5** 28.6%***	7 3.5 33.3%	8 14.1 38.1%	0 1.6 0.0%	0 0.3 0.0%	21 7.7%
2 139- 130	6 2.8 15.0%	16 6.7 40.0%	18 26.8 45.0%	0 3.1 0.0%	0 0.6 0.0%	40 14.7%
3 129- 120	2 5.2 2.7%	19 12.5 25.7%	53 59.6 71.6%	0 5.7 0.0%	0 1.1 0.0%	74 27.1%
4 119- 110	4 4.2 6.6%	1 10.3 1.6%	48 40.9 78.7%	8 4.7 13.1%	0 0.9 0.0%	61 22.3%
5 109- 100	1 4.0 1.8%	3 9.6 5.3%	40 38.2 70.2%	11 4.4 19.3%	2 0.8 3.5%	57 20.9%
6 99- 90	0 1.2 0.0%	0 2.9 0.0%	13 11.4 76.5%	2 1.3 11.8%	2 0.2 11.8%	17 6.2%
7 89 and below	0 0.2 0.0%	0 0.5 0.0%	3 2.0 100.0%	0 0.2 0.0%	0 0 0.0%	3 1.1%

* actual number
** expected value
*** row percentage

Table 6

1985 I.Q. Test
Population = 249

Test Range	1 Lose 20+	2 Lose 19-11	3 Lose 10 or gain 10	4 Gain 11-19	5 Gain 20+	Total
1 150+ 140	4* 0.5** 44.4%***	3 0.7 33.3%	2 6.4 22.2%	0 1.3 0.0%	0 0.2 0.0%	9 3.6%
2 139- 130	4 0.9 22.2%	6 1.4 33.3%	8 12.7 44.4%	0 2.5 0.0%	0 0.4 0.0%	18 7.2%
3 129- 120	2 3.3 3.1%	4 4.9 6.3%	55 45.2 85.9%	3 9.0 4.7%	0 1.5 0.0%	64 25.7%
4 119- 110	1 3.8 1.4%	3 5.5 4.2%	58 50.9 80.6%	9 10.1 12.7%	1 1.7 1.4%	72 28.9%
5 109- 100	2 2.8 3.8%	2 4.0 3.8%	35 37.5 66.0%	12 7.4 22.6%	2 1.3 3.8%	53 21.3%
6 99- 90	0 1.3 0.0%	1 1.9 4.0%	13 17.7 52.0%	9 3.5 36.0%	2 0.6 8.0%	25 10.0%
7 89 and below	0 0.4 0.0%	0 0.6 0.0%	5 5.7 62.5%	2 1.1 25.0%	1 0.2 12.5%	8 3.2%

* actual number
** expected value
*** row percentage

Table 7

1986 I.Q. Test
Population = 236

Test Range	1 Lose 20+	2 Lose 19-11	3 Lose 10 or gain 10	4 Gain 11-19	5 Gain 20+	Total
1 150+ 140	5* 0.4** 55.6%***	3 0.9 33.3%	1 6.0 11.1%	0 1.3 0.0%	0 0.4 0.0%	9 3.8%
2 139- 130	2 1.0 9.5%	5 2.0 23.8%	14 14.0 66.7%	0 3.0 0.0%	0 1.0 0.0%	21 8.9%
3 129- 120	3 2.1 6.5	5 4.5 10.9%	36 30.6 78.3%	2 6.6 4.3%	0 2.1 0.0%	46 19.5%
4 119- 110	0 2.8 0.0%	5 5.9 8.2%	47 40.6 77.0%	7 8.8 11.5%	2 2.8 3.3%	61 25.8%
5 109- 100	1 2.8 1.6%	2 5.9 3.3%	42 40.6 68.9%	13 8.8 21.3%	3 2.8 4.9%	61 25.8%
6 99- 90	0 1.4 0.0%	3 2.9 10.0%	13 20.0 43.3%	9 4.3 30.0%	5 1.4 16.7%	30 12.7%
7 89 and below	0 0.4 0.0%	0 0.8 0.0%	4 5.3 50.0%	3 1.2 37.5%	1 0.4 12.5%	8 3.4%

* actual number
** expected value
*** row percentage

Statistically the "standard error of measurement is larger for high scores than for average scores because the number of items on which higher scores are based is fewer than the number of items on which average scores are based."[1] When items are selected for inclusion on a group test, the majority of the items included must be correctly identified by most of the group to be tested. By imposing this restraint in the development of the group test, differentiation at the upper end is sacrificed. Table 8 demonstrates how the very design of the Cognitive Abilities Test is such that one item, or being three months younger (even though time in school may be the same), can be the difference between inclusion into the gifted program or exclusion from it.

Table 8

Raw Score Conversion to SAS

Raw Score	Standard Age Score (6-9)	(7-0)	(7-3)	(7-6)	(7-9)
76	150+	150+	150	148	147
75	149	146	143	141	137
74	143	140	137	134	131
73	139	136	133	130	127
72	134	131	131	128	125
71	130	127	124	121	118
70	126	123	120	117	114

[1] Linda Silverman, "Giftedness, Intelligence and the New Stanford-Binet," Roeper Review 8 (Feb. 1986): 169.

It should be noted that a raw score of 76-74 is ranked in the 99%, 73 is 98%, 72 is 96%, 71 is 94%, and 70 is 92% according to the examiner's manual of Level 1. However, inclusion in a gifted and talented program in Iowa is limited to the top 3 percent and as Tables 2-7 showed, the top 3 percent was always in the 140 and above range.

Discussion

This study has shown that the scores from the first grade test are unreliable indicators of fourth grade scores at the upper range. This paper does not address the definition of giftedness. The issue that is explored in this study is the reliability, or repeatability, of the Cognitive Abilities Test in identifying the top students of a school district. Whatever definition of giftedness is used, the fact is that Cognitive Abilities or similar tests are used by the majority of school districts as a tool for identification and thus placement in a gifted and talented program.

Due to the high reliability rating of the Cognitive Abilities Test, it is used extensively.[1] As this study showed, it is very reliable/repeatable at the mid-range band of scores. It is easy to assume that this reliability/ repeatability holds true throughout the whole range of

---

[1] Mitchel.

scores. However, this study has shown that to be an erroneous assumption. It is at the upper ranges, the very range where scores are being used to place and to eliminate students, that this test lacks reliability. This lack of reliability/repeatability in the upper range bands of scores is evident in all five years of this study. This lack of reliability/repeatability in the upper range bands of scores is present in every one of the eight schools, in every year of this study.

The 140 to 150+ range comprised 6 percent of the population in 1982, 8 percent in 1983, 7.7 percent in 1984, 3.6 percent in 1985, and 3.8 percent in 1986 for an overall 5.9 percent in the summary table. In Iowa, gifted and talented programs are to be composed of students in the top 3 percent of the population. Thus, in each and every year for five consecutive years this 3 percent is in the 140-150+ range. Yet, as Tables 2-7 clearly show, losses of 11 or more points in 90 percent of the scores in 1982, 85 percent in 1983, 62 percent in 1984, 78 percent in 1985, and 89 percent in 1986 with the five year summary showing 80 percent loss in this range. Never is the retention or normal band of plus or minus 10 points ever maintained in even close to half of the students' scores in this range. This brings into question the use of scores of this group test for differentiation such as placement in or exclusion from a gifted program at the very range where the repeatability of the test is shown to be so poor.

In order to obtain a 140 or above SAS a student can only miss one to three items. The fact that a seven year old can miss three items and a student only three months older can miss two items while a seven year nine month old student is allowed only one miss, gives a definite advantage to the younger student even though time spent in school may be the same for all the students.

An example of one of the items on the first grade test is a row containing the picture of a cow, a horse, a pig, and a bird. The student is directed to mark the item that is different. According to the manual the bird is the different animal. However, a child had correctly identified the bird and then changed the answer to the pig. There is a container of food in front of the pig, so the pig is the only animal pictured that is eating. Another example of an item is one where the student is directed to identify the meter; there are four things pictured, one being a parking meter which, according to the manual, is the correct choice. Many youngsters living in a suburb have never seen a parking meter and identify the object that appears to be a ruler--the meter stick. These examples cited are in no way meant to be a criticism of the test per se, but are intended to point out that a student may be penalized with an incorrect answer, when upon examination of the thinking process of the student, the choice made does not appear to be without merit. Certainly we should recognize the student who looks at things perhaps from a different perspective. The discovery of the double helix of the DNA molecule by a man with an I.Q. of 115 would seem to represent someone whose thinking was different than most.

Since the I.Q. test was originally designed to identify those who would not benefit from schooling and today group tests are not allowed to be used to identify someone for placement in a remedial program it is ironic and perhaps inappropriate that they are used for placement or exclusion in a gifted program. An analysis of the Standard Age Score shows the plus or minus ten point category has the greatest showing in the mid-range scores (119-100). Thus the change occurs within the predicted range. However, it is the upper range (140+) scores that are used for identification for a gifted program and this is the very spot where the scores lose their repeatability and therefore their reliability.

Gardner's description of intellectual competence incorporates many of the abilities cited by those previously mentioned and he calls for "a set of skills of problem solving--enabling the individual to resolve genuine problems or difficulties that he or she encounters and, when appropriate, to create an effective product--and must also entail the potential for finding or creating problems--thereby laying the groundwork for acquisition of new knowledge."[1]

Certainly gifted and talented students should possess the ability to solve new problems, acquire new knowledge, and when appropriate to create an effective product. These are not traits identified through group I.Q. tests.

This study has demonstrated that there is inconsistency in the ability of the Cognitive Abilities Test to identify the gifted and talented. Educators should be urged to seek more reliable means for identification.

"An individual can lose his entire frontal lobes, in the process becoming a radically different person, unable to display any initiative or to solve new problems--and yet may continue to exhibit an IQ close to genius level."[2]

---

[1] Gardner, p. 60.

[2] Gardner, p. 18.

BIBLIOGRAPHY

Books

Cronbach, L. J. Essentials of Psychological Testing. New York: Harper & Row, 1970.

Dearborn, W. F., and J. W. M. Tothney. Predicting the Child's Development. Cambridge: Science Art, 1941.

Gallagher, James J. Teaching the Gifted Child. Boston: Allyn & Bacon, 1975.

Gardner, Howard. Frames of Mind. New York: Basic Books, 1983.

Goldstein, Gerald, and Michael Hersen. Handbook of Psychological Assessment. New York: Pergamon, 1984.

Guilford, J. P. The Nature of Human Intelligence. New York: McGraw-Hill, 1967.

Jensen, Arthur R. Bias in Mental Testing. New York: Free Press, 1980.

Mitchel, James V., Jr., ed. 9th Mental Measurement Yearbook. Lincoln: Univ. of Nebraska Press, 1985.

Sattler, Jerome. Assessment of Children's Intelligence and Special Abilities. Boston: Allyn & Bacon, 1982.

Sharp, Evelyn. The IQ Cult. Toronto: Longman Canada Ltd., 1972.

Sternberg, R. J. Handbook of Human Intelligence. New York: Cambridge Univ. Press, 1982.

Tannenbaum, Abraham J. Gifted Children. New York: Macmillan, 1983.

Terman, L. M. Genetic Studies of Genius: Mental and Physical Traits of a Thousand Gifted Children. Calif.: Stanford Univ. Press, 1926.

Thorndike, Robert L., and Elizabeth Hagen. Cognitive Abilities Test: Examiners Manual, Form 3. Atlanta: Houghton Mifflin, 1979.

Vernon, Philip E. Intelligence: Heredity and Environment. Alberta: Univ. of Calgary, 1977.

Whitmore, J. R. Giftedness, Conflict, and Underachievement. Boston, Allyn & Bacon, 1980.

Periodicals

Alvino, James, Rebecca McDonnel, and Susanne Richert. "National Survey of Identification Practices in Gifted and Talented Education." Exceptional Children 48 (Oct. 1981): 124-32.

Borland, James. "I.Q. Tests: Throwing Out the Bathwater, Saving the Baby." Roeper Review 8 (Fe. 1986): 163-67.

Eysenck, H. J. "Recent Advances in the Theory and Measurement of Intelligence." Early Child Development Care 15 (1984): 97-115.

Feldhusan, John F., and Steven M. Hoover. "A Conception of Giftedness: Intelligence, Self Concept and Motivation." Roeper Review 8 (Feb. 1986): 140-42.

Gagnel, Francoys. "Giftedness and Talent: Reexamination of the Definitions." Gifted Child Quarterly 29 (Summer 1985): 103-11.

Hatch, Thomas C., and Howard Gardner. "From Testing Intelligence to Assessing Competences: A Pluralistic View of Intellect." Roeper Review 8 (Feb. 1986): 147-50.

Honzik, M. P., J. W. MacFarlane, and L. Allen. "The Stability of Mental Test Performance Between 2 and 18 Years." Journal of Experimental Education 17 (1948): 309-24.

Kaufman, Alan S., and Patti L. Harrison. "Intelligence Tests and Gifted Assessment: What Are the Positives?" Roeper Review 8 (Feb. 1986): 154-49.

Marland, S. P. Education of the Gifted and Talented. Report to the Congress of the United States by the U.S. Commissioner of Education, Vol. 1. Washington, D.C.: U.S. Government Printing Office, Aug. 1971.

Martinson, R. A., and L. M. Lessinger. "Problems in Identification of Intellectually Gifted Pupils." Exceptional Children 26 (1960): 227-42.

Moley, B. E., F. A. Olson, T. Haleves, and J. Flavell. "Production Deficiencies in Young Children's Clustered Recall." Developmental Psychology 1 (Jan. 1969): 26-34.

Pegnato, C. W., and J. W. Birch. "Locating Gifted Children in Junior High Schools Comparison Methods." Exceptional Children 25 (Mar. 1959): 300-304.

Renzulli, Joseph S., and Marcia A. B. Delcourt. "The Legacy and Logic of Research on the Identification of Gifted Persons." Gifted Child Quarterly 30 (Winter 1986): 20-23.

Robinson, Nancy M., and Diana L. Chamrad. "Appropriate Uses of Intelligence Tests with Gifted Children." Roeper Review 8 (Feb. 1986): 160-63.

Scherer, M. "How Many Ways is a Child Intelligent?" Instructor 94 (Jan. 1985): 32-35.

Silverman, Linda. "Giftedness, Intelligence and the New Stanford-Binet." Roeper Review 8 (Feb. 1986): 168-71.

Sternberg, Robert J. "Identifying the Gifted Through IQ: Why a Little Bit of Knowledge is a Dangerous Thing." Roeper Review 8 (Feb. 1986): 143-47.

---. "Testing Intelligence without I.Q. Tests." Phi Delta Kappan 65 (June 1984): 694-700.

Sternberg, Robert J., and Janet Davidson. "The Mind of the Puzzler." Psychology Today (June 1982): 37-44.

Sternberg, Robert J., and Janet S. Powel. "Comprehending Verbal Comprehension." American Psychologist 38 (1983): 878-93.

Sternberg, Robert J., and R. K. Wagner. "Alternative Conceptions of Intelligence and Their Implication for Education." Review of Educational Research 543 (1984): 179-223.

Torrance, E. P. "The Role of Creativity in Identification of the Gifted and Talented." Gifted Child Quarterly 28 (1984): 153-56.

Treffinger, Donald J., and Joseph Renzulli. "Giftedness as Potential for Creative Productivity: Transcending IQ Scores." Roeper Review 8 (Feb. 1986): 150-54.

Wolf, Lee. Telephone interview. 21 July 1986.